

## Voices at the Table: Balancing the Needs and Wants of Program Stakeholders to Design a Value-added Writing Assessment Plan

Terry Myers Zawacki, E. Shelley Reid, Ying Zhou and Sarah E. Baker, George Mason University

**Abstract:** The WAC director, composition director, director of Institutional Assessment (OIA), and WAC/OIA liaison describe the programmatic "needs" and "wants" they balanced in the plan they collaboratively designed to respond to a state mandate for "value-added" writing assessment. To satisfy this mandate, as they explain, they carried out an assessment of research-based essays from the first-year composition (FYC) course that mirrored the discipline-focused, course-embedded, and workshop-based assessment process George Mason has been implementing successfully in upper-division writing-intensive (WI) courses since 2002. Just as the writing-in-the-disciplines (WID) assessment data have been informing WAC efforts, the FYC process was designed to provide data to inform programmatic efforts in curriculum and faculty development. The authors begin with background on their model program of WID-based assessment and then discuss the steps they took to develop a value-added proposal that included well-defined learning outcomes for both introductory composition and the WI courses and that was reliable, valid, cost-effective, and sustainable. With this backdrop, they each describe and reflect on their different perspectives on the FYC implementation process, program goals, and the assessment outcomes. They conclude by offering their collective views on how the overall process has fostered the kinds of cross-disciplinary conversations that help them sustain and enhance their programs, and how these conversations model the spirit of negotiation and cooperation that has likewise sustained the culture of writing at Mason.

In 2007, in the context of a national dialogue that was increasingly focused on accountability and assessment of student learning outcomes, the Virginia State Council of Higher Education (SCHEV) appointed a [task force](#) to revise its competency-based guidelines on student learning outcomes assessment. The new guidelines, approved in January 2008, required institutions to conduct value-added assessment that embodied the following concept: "Value-added assessment measures indicate progress, or lack thereof, as a consequence of the student's institutional experience" (p. 7). This is the voice in the background, weighty with legislative authority, that brought the four of us at George Mason University to the planning table in spring 2008 to discuss ways we might demonstrate that our writing instruction itself was adding value to students' overall educational experience—while

---

### *Across the Disciplines*

*A Journal of Language, Learning and Academic Writing*

DOI: <https://doi.org/10.37514/ATD-J.2009.6.1.03>

wac.colostate.edu/atd

ISSN 554-8244

*Across the Disciplines* is an open-access, peer-review scholarly journal published on the WAC Clearinghouse and supported by [Colorado State University](#) and [Georgia Southern University](#). Articles are published under a [Creative Commons BY-NC-ND license](#) (Attribution-NonCommercial-NoDerivs) ISSN 1554-8244. Copyright © 1997-2017 The WAC Clearinghouse and/or the site's authors, developers, and contributors. Some material is used with permission.

still retaining the discipline-focused and workshop-based assessment process we'd been successfully implementing since 2002. As the voices in the foreground, we resisted solutions that had been proposed in initial discussions about the new mandate: for instance, that we administer a timed writing test to all incoming students in our introductory writing course or collect an essay written in the first weeks of the course. We had no interest in endorsing measures we'd already rejected as inconsistent with Mason's overall assessment philosophy when we developed our current course-embedded process for assessing students' competence as writers in their majors. Moreover, we were committed to a value-added approach that would allow for a pre-assessment to be embedded in the first-year composition (FYC) course, thereby providing data to inform programmatic efforts in curriculum and faculty development, just as the writing-in-the-disciplines (WID) data have been informing WAC efforts.

Bringing our first-year composition course into a WID-based writing assessment process posed both opportunities and challenges, almost as though we were adding another discipline to the culture of writing we pride ourselves on having established at Mason. While the writing center and our required [Advanced Composition](#) classes (focused variously on Humanities, Social Science, Business, Natural Science, and Technology) very deliberately serve students who write within and across disciplines, [English 101](#) has traditionally identified more with a national concept of FYC (exemplified by the [WPA Outcomes Statement](#)) than with our local writing culture. Indeed, English 101, like some FYC courses elsewhere, might aptly be described as a course that teaches students that "good writing is good writing," without an explicit focus on how that education might transfer into disciplinary contexts. Our conversations throughout the value-added planning and implementation process, then, cannot be characterized as "the writing people vs. the assessment people," but involved a much richer, more nuanced, multi-vocal approach, with each of our voices coming to the front or receding as we worked through a collaborative process of discovery.<sup>[1]</sup>

In this article we represent those voices and the inductive discoveries that emerged from our earliest conversations through our development of the protocols, our implementation of the new assessment processes, and our reflections on what we've gained and hope yet to gain from our participation in this process. An initial section giving Terry's perspective on the WID-based assessments already in place thus serves as a backdrop to Ying's discussion of developing value-added approaches that originated in the Office of Institutional Assessment (OIA), and to Shelley's and Sarah's analysis of how the new protocols worked "on the ground" in the composition program, and their reflections on how participants in this kind of cross-disciplinary process need to attend to relationship-building as well as to the actual rubrics and events. In the final section, we offer our collective analysis of the discoveries and gains we made as we negotiated our respective "needs" and "wants," hoping to illustrate how such a process, with its surprises and compromises, can yet produce a range of satisfactory results and foster sustainable cross-disciplinary conversations about student competencies, writing curricula, faculty development, and larger university writing cultures.

## A History of Mason's WID-based Assessment

### Terry

When Ying first told me about the SCHEV's most recent requirement for value-added assessment, I confess my initial response was to say that I was not interested in negotiating any aspect of the WID workshop process we'd been implementing successfully for the past seven years in order to satisfy yet another mandate. My next, more measured response was to suggest that we bring Shelley to the table to brainstorm a method for gathering "pre" data that would mirror our WID assessment methods and the "post" data we'd been gathering on students' writing competence. Sarah was

already a knowledgeable participant in the conversation, as she had been assisting with both writing assessment and WAC for the past year; the value-added mandate had, in fact, provided the catalyst for formalizing her jointly funded position as a liaison for OIA, WAC, and departments engaged in writing assessment. In her liaison role, Sarah is charged with the often unenviable task of representing the "needs" of the assessment office to faculty constituencies who, for the most part, don't think they need to assess students' writing competence nor want to spend time doing so. In her WAC role, however, she assists me in the much more satisfying and enlightening process—for all involved—of conducting the departmental writing assessment workshops I describe next.

While 2002 marked the beginning of the workshops, WAC and OIA have had a close working relationship for many years based on a mutual understanding of the purpose of assessment as a way to improve, not just to prove, student learning and program effectiveness. It is this larger view of assessment as being adaptable to serve educational goals rather than only serving accountability that led to the formation of the Writing Assessment Group, composed of faculty representatives from across the colleges and charged, initially, with helping to develop an assessment plan that would also provide WAC faculty development opportunities. That we had the latitude to develop our own plan, as well as to define written competence for our students, speaks also to the value SCHEV itself placed on having the state institutions design assessment processes that reflected their particular contexts and interests. The plan our group submitted defined writing competence very generally as the ability to use writing to learn and to express knowledge, but we also explained that, because disciplines have distinct and different goals for student writers, we intended to focus our assessment on student writing in the majors, embedding the process in required WI courses using papers selected by faculty and assessed with a rubric faculty developed together in a holistic scoring workshop. With its focus on student papers in the major, the workshop was also designed to allow for a wider discussion of teachers' expectations for student writers and how these are conveyed to students through assignments, comments on papers, grades, and grading criteria. Over the years, the Writing Assessment Group has continued to meet in an advisory capacity, and, when called upon, assists me and Sarah in implementing the process in departments.

Our WID assessment process and the ways the resulting data inform WAC efforts have been written about in other venues, and Mason is featured as a model program on the WPA website's "[Assessment Gallery and Resources](#)," so I won't go into detail here other than to say that our process reflects our strong WAC culture and the principles that guide us, i.e., that students learn to become competent writers when they have frequent opportunities to practice the genres and conventions typical of their majors along with opportunities to revise based on meaningful feedback from teachers in and across the disciplines, in this way benefitting by writing for multiple and diverse audiences. In addition, we are enacting principles that have been articulated by writing assessment scholars for decades, from Charles Cooper on the holistic evaluation of writing and Barbara Walvoord on WID-focused assessment and course-embedded procedures to Brian Huot on meaningful assessment. Huot argues that the best assessment is site-based and locally controlled with questions and measures developed by those who have a stake in the process and the results; led by writing professionals; grounded in theory; and conducted with a conscious awareness of the beliefs and assumptions underlying our actions. For Huot, writing assessment is a "social action" in that it can help us shape instruction that considers the writing needs of all students.

These principles are closely aligned with an assessment philosophy Karen Gentemann, the Associate Provost for Institutional Effectiveness, describes in a forthcoming article she and I co-authored:

..."assessment" is not the equivalent of "testing," but is rather a philosophy about education, albeit accompanied by an emerging consensus of what constitutes good

methodology and best practice. The philosophy, simply stated, is that student learning is the purpose of teaching and that much of student learning can be demonstrated, and, further, if a good assessment is conducted, corrections or changes can be made to enhance the learning experience for students.

The concept that faculty own the curriculum is central to assessment practice, Gentemann explains, yet she is also careful to point out that individual faculty do not "stand alone" in their classrooms. Rather they must join with others in the program to "establish coherence in the curriculum by agreeing upon the contribution of each part and sharing a sense of direction and purpose for the student and the learning experiences." When faculty share goals for student learning, they are also more inclined to feel they have a stake in the results of their assessments because the information will help them make changes in their courses and in the curriculum. "Thus faculty must be involved, at some level, in developing and participating in the process," Gentemann argues, "and, whatever their role, they must be vested in knowing the results of the assessment."

Working from these principles for meaningful assessment and our shared interest in placing responsibility for assessment with faculty in the majors, Gentemann and I, in conversation with the Writing Assessment Group, had devised the workshop process I described earlier, which successfully combines faculty development with the required scoring and reporting. One of the great values of these workshops is that they bring together faculty within a department who may teach with writing yet may never have sat down with their colleagues for a pedagogically focused discussion of their expectations for student writers related to the genres and conventions that characterize good writing in their disciplines. Most often, then, we encounter an initial "good writing is good writing" attitude (along with questions about why we can't just use a generic rubric for scoring), an attitude which quickly changes once faculty begin discussing the traits they value in the four sample papers selected to begin the workshop scoring process. Further, over the years we've been conducting workshops, many faculty participants have revised assignments, assignment sequences, and their commenting practices based on insights and information gained from discussing the sample papers with their colleagues in the process of generating a rubric for assessing their students' writing in a curricular context.

Change has also occurred at the department level. One department, for example, decided to change the designated WI course because its focus and the writing being assigned were not appropriate to the larger writing outcomes the faculty wanted students to achieve, the latter a topic that was brought up and discussed during the course of the workshop. Many departments have circulated the assessment rubric to their faculty, advising them to use the rubric as a guideline and/or to adapt it to fit their own assignments and expectations for writers in their courses. Some departments have created [online writing guides](#) to help students understand the ways of knowing and doing, to echo Michael Carter, that are most characteristic of their disciplines and subdisciplines. For all of these reasons and more, we were satisfied that the assessment workshops were achieving the results we'd intended and were helping faculty to improve their courses and curriculum at the same time as they were providing results to SCHEV.

As the WAC program stakeholder, then, I was adamant about retaining our WID-based workshop process when the four of us came to the table for the next round of planning to accommodate SCHEV's value-added requirement. All of us, as I noted earlier, were also committed to developing an approach that would mirror this process in its methods and the kinds of useful data it promised to yield for the composition program. While we were all in agreement on these points, however, the finer "needs" and "wants" details remained to be negotiated. In the next sections, Ying, Shelley, and Sarah explain what was at stake for them in the negotiations, the new plan that emerged, the "on-the-ground"

implementation, and the discoveries they've made in the process. It's worth noting here that, in her liaison role, Sarah played a key role in developing and implementing the FYC assessment plans; with her experience as an assistant to the Writing Assessment Group and the WAC Committee for three years and as an adjunct in the composition program, she has a good understanding of how to interpret and communicate our various programmatic interests and was able to balance these with skill and insight to help us achieve shared program goals. While, at Mason, Sarah may be uniquely qualified by experience and position to play this role, we believe that our example demonstrates that any assessment process can be designed to deliberately foster collaboration and participation if the key program stakeholders are brought to the table for conversations like those we'll describe.

## Balancing Stakeholder "Needs" and "Wants" for a Proposal to SCHEV

### Ying

In early January 2008, when SCHEV officially asked institutions to submit value-added assessment proposals for two competency areas (written communication and scientific reasoning) by the first of March, I realized that I would need to collaborate closely with relevant faculty members in those areas, but especially with composition faculty, for many of the reasons Terry explains above. While pondering different options for assessing written communication, I firmly believed that the "natural" place to conduct a pre-assessment was first-year composition, our English 101 course. Further, given SCHEV's short reporting deadline and the fact that most Mason students take their first composition course in the fall semester, it seemed to me that our only window for data collection was fall 2008. To meet SCHEV's guidelines, our plan needed to include the following features:

1. *Well-defined learning outcomes for both English 101 and the WI courses.* While English 101 has a common set of learning outcomes based on WPA outcomes, our WI courses, where our post-assessment occurs, define outcomes in the context of the discipline. What we needed, then, were learning outcomes at the WAC program level that clearly articulated what our students would be able to demonstrate after going through their composition and WI courses.
2. *Reliability and validity.* Reliability was not a concern because at the WAC assessment workshops faculty raters not only develop the scoring rubrics together but also practice scoring using student papers. The training process inherent to the workshops improves inter-rater reliability. However, I was concerned that validity could be an issue because, although the writing is authentic, the writing assignments in English 101 and WI courses come in a variety of forms, which would have an impact on our ability to assess students' competence. For example, by examining annotated bibliographies from ENGL 101 and research papers from WI courses, we may not be able to conclude that our writing program has added value to student growth.
3. *A good sampling strategy.* We needed to collect sufficient and representative writing samples. Sampling is technically easy, but faculty participation and compliance can be an issue when they believe the process is an excessive and cumbersome requirement rather than an opportunity for development. A robust data collection strategy would call for mandatory participation of all sections of English 101 and a random sample of enrolled students.
4. *Cost effectiveness.* SCHEV's new mandate did not come with any funding, and, while tenure-track faculty had been the main participants in the WID-assessment workshop and, as such, had not been compensated, the English 101 instructors, predominantly adjuncts, would need to be

paid for the additional work they put into assessment. I knew we could offer only a very small stipend for any voluntary raters who were not tenure-track faculty.

5. *Sustainability and efficiency.* I shared with the others the belief that any plan we developed needed to be sustainable and focused on improving, not only on proving. But, with little to no funding, I initially thought a timed test administered in English 101 might be the most efficient option.

With those concerns in mind, I called a meeting with Terry, Shelley, and Sarah at the end of January, five weeks before the proposal was due. It was one of the best meetings I have had with faculty. They did not ask the most frequent questions I get: "Why this, why now, and why us?" Nor did they ask the second most frequent question: "What's the minimum we can do to satisfy SCHEV?" Instead, they saw the new mandate as another opportunity for faculty development and curricular improvement. As we discussed possible options, we finally decided to build on Mason's tradition of faculty-led assessment of writing in the disciplines and embed the pre-assessment in English 101. No writing test would be conducted at the beginning of the course; instead, a research-based essay, assigned in almost all sections of English 101, would be used for assessment.

In essence, our value-added assessment approach measures students' writing competence at two course levels, a lower-division FYC course and upper-division WI courses, using representative samples of research-based papers assigned in the class. Faculty, including FYC faculty, develop writing rubrics and rate sample papers using the same process Terry described earlier. The only additional requirement is that they now must include in their writing rubrics a benchmark "overall writing competence" that allows us to compare competence levels between first-year and upper-division students to find out whether Mason's curriculum adds value to their growth in writing. Our approach is different from that of some Virginia institutions, as I will describe later, but it aligns with SCHEV's definition.

The four of us left the meeting table with a clear assessment process identified, but with details still to be worked out. At the proposal stage, the easier part for me was to describe the value-added analytical approach using assessment language, to justify it in the context of Mason's curriculum and student population, to explain data collection techniques and measurement strategies, and to estimate the cost. Harder, however, was defining general learning outcomes for the composition and WAC programs in order to specify general criteria for writing competence that SCHEV had explicitly asked for. I began by going through all the departmental rubrics from the writing assessment workshops from 2002-2007 to identify the most frequently occurring criteria. Next I drafted a [writing assessment checklist](#), which, I thought, would serve two important purposes: 1) it would show SCHEV the criteria we use to measure writing competence, and 2) it would help Mason faculty develop their own writing rubrics by providing a pool of commonly used criteria. Terry, Shelley, and Sarah supported the first purpose, but were strongly against using the checklist in any faculty scoring workshops. They argued that presenting faculty with a ready-made list of criteria from which to choose went against the organic process of developing a discipline-based rubric that was at the heart of the assessment workshop model. Although we did update and refine the checklist, it ultimately was used just for SCHEV. One important outcome of the discussion, however, was that Terry was able to use the general criteria checklist as the foundation for a set of commonly held learning goals for student writers across disciplines that she included as part of the WAC mission statement and goals already articulated on the [WAC program site](#).

Another issue to be worked out concerned how we should define the benchmark for "overall writing competence." Some individual WID rubrics had included an "overall writing" category that defined

three levels of competence, "more than satisfactory," "satisfactory," and "less than satisfactory." These three levels, although sufficient for rating papers from WI courses, were not sufficient to demonstrate student growth from freshman year to junior year. For benchmarking, we would have to define at least four levels of competence, with one level being "emerging competence," to allow English 101 students to be judged as "satisfactory" at the level of FYC but less than satisfactory for the WI level. After much discussion, the four levels were fixed upon as: highly competent, competent, emerging competence, and not competent college-level writing, with the criteria and standards for each level defined in detail. We agreed that for both pre- and post-assessment, faculty would continue to develop and use their own scoring rubrics, but they would use these four levels for an overall rating. We expected that most FYC students would fall into the "emerging" category and that most students taking WI courses would be judged "competent."

With a plan in place and [the proposal](#) sent off to SCHEV, I prepared myself for the reviewers' feedback. Although the state allows flexibility in assessment plans, it has a very rigorous review process of the proposals. A peer review of our plan was completed in May 2008 by assessment professionals from peer institutions in the state. As I expected, my assessment colleagues interpreted SCHEV's "value-added" definition differently, proposed different methods based on their institutional contexts, and had different views of our proposal. One reviewer found our plan appropriate, and the other recommended we improve our method by standardizing writing prompts. In our response, we acknowledged the reviewers' concerns but argued—successfully, as it turned out—that our plan would be the least intrusive to faculty teaching, and the data we collected would be most useful for the composition program and individual academic programs. We also reiterated the high value our university places on the culture of writing we've created over the decades, with assessment playing a key role in sustaining that culture by enabling the conversations about writing that Terry has described. After going through the review process, two outcomes were clear to me: 1) carrying out a value-added writing assessment is doable, but a real challenge to every institution; and 2) the implementation process will not be as clean as stated in any institutions' proposals.

## Bringing First-Year Composition to the Table

### Shelley

I had known for years that the composition program at Mason needed assessment, but building a workable assessment program from scratch seemed a daunting task. Likewise, I knew that any curricular revision that might bring our first-year course into closer alignment with other elements of our campus-wide writing programs would be difficult. Mason's faculty and courses are about as decentralized as is possible in a responsible FYC program. English 101 (like its sheltered-ESL equivalent, English 100) is taught in large part by our newest adjunct faculty and our graduate teaching assistants: faculty members who overall have the least experience, the least connection to our program, and the fewest resources of any in our department. Additionally, we do not stipulate a textbook or a curriculum; the course is governed instead by recently revised learning outcomes that draw on the WPA Outcomes Statement, and our oversight of how more than teachers in 100 sections a year implement these outcomes is necessarily thin. Consistency and growth have to come by agreement and opportunity: through faculty development, curriculum resources, and gentle suasion. We thus needed three things that only large-scale outcomes assessment could provide:

- general knowledge about instructor and student practice;
- faculty development work based on an organized, collaborative review of student writing; and
- a way to establish more benchmarks by which new and continuing faculty can orient their work.

Certainly the best way to approach a faculty community whose members thrive on individual control over their teaching is to build the assessment program from within. Ideally, too, the most useful knowledge of instructor and student practices would come from a portfolio review to assess students' abilities as readers, writers, and revisers in multiple genres. And the most effective faculty development would involve large numbers of faculty participating in multiple facets of the process, from design to rating to planning for subsequent recommendations. As Bob Broad explains, there is much to be learned in the conversation among faculty about how we value student writing. Benchmarks that result from such a process more closely match and engender good local practices, working as magnets to draw faculty together rather than fences to keep them from straying. So I entered this conversation wanting the richest, most home-grown assessment protocols that we could develop—but needing the external mandate, funding, and assistance that the SCHEV-initiated assessment could offer. In theory, these demands can be balanced even in a large, dispersed program like Mason's; in practice, we found that our successes were uneven and contingent, but that our emphasis on collaboration was crucial in getting as far as we did.

At the start, the collaborative approach we took in response to the SCHEV mandate let me find a middle ground between assessing and being assessed. I felt at each step that I was both learning and teaching, that my suggestions were being listened to and improved upon, and that I not only could but was being asked to use the process to reach goals that I had for the composition program, within the resources available to the program and my colleagues. Moreover, our decision to situate the composition assessment within the overall culture of writing at Mason helped me feel that I was contributing to an ongoing process rather than being singled out for additional work. Quite often, in my experience, the "do as little as possible" assessment model focuses on a single moment in time or a very short arc of student learning in order to very firmly and expediently "close the loop"; if faculty then feel as though they have invested a lot of time for very little return, it is difficult to contradict that response. In this case, though, the four of us began our discussion by seeking to address how the extant upper-division assessment, with its emphasis on discipline-appropriate writing, would relate to FYC assessment. The "emerging competence" rating emphasized how our English 101 assessment efforts would fit into a larger continuum of student learning. The process we designed is, as we describe below, not just one that we could and might want to replicate later, but is also itself replicating—and thus sustaining—assessment that was already in place at the university. The composition program thus took a step toward becoming part of a community of assessment as well as a community of writing.

## Making it Work: The Assessment Process, Goals, and Outcomes

### Shelley and Sarah

At the beginning of the fall semester, we sent all English 101 instructors a one-page outline detailing the assessment plan. We had designed our collection process to generate as little additional work for faculty and provide as much motivation and information for them as possible, as well as to keep communications flowing. Instructors were asked to collect a clean final copy (paper or electronic) of a research-based assignment from four randomly selected students in each section, a larger-than-necessary sample to allow for any problems with compliance. In November, Sarah emailed instructors with the names of the students from whom they were to collect papers and the request for the assignment prompt; students were also given a brief letter of explanation. Sarah chose to email each instructor separately because she felt that it might improve compliance: from her experience tutoring in the Mason writing center and teaching in the composition program, she personally knew half the instructors. By not being a faceless administrator, Sarah not only made it



easier to respond to individual queries or concerns but also kept the assessment conversation going in virtual and actual hallways. The final collection rate was 71.7%, with papers collected from 79 out of 85 sections (93%) of English 100 and 101—an instructor-compliance rate actually a little higher than our current rate for submitting syllabi to the department files. It's likely that this rate would have been lower had we asked for more materials (drafts, reflective writing, instructor input), or had we not been able to use both local/personal appeals and the specter of the state requirement to induce participation.

An initial plan for a set of collaborative rubric-development workshops with English 101 faculty dwindled to a single meeting amid time pressures and other factors. Our workshop began with a discussion of the four rating-levels that our committee had designed as the common denominator for developmental progress. Because we were aiming to rate FYC papers in the context of the broader community of writing, we looked at some of the upper-division WID rubrics currently in circulation. Then the 10 of us—Sarah, Shelley, and Terry, along with three full-time non-tenure-track faculty, three adjunct faculty, and one TA—began to generate our own criteria for evaluating student writing. Nearly every criterion we generated provoked intense discussion: What did we mean by it? Should it be placed in another category? Was it more or less important in an overall rating of a student essay? When we generated a draft FYC rubric and used it to assess two sample papers, the conversations deepened: was "analyzes and synthesizes source material" better as an element of the category "Use of Sources" or of the category "Contribution to Conversation"? Was there a level of grammatical error that one could imagine would automatically lower an essay's ranking? Was Student X's third paragraph evidence of synthesis or just of quilting quotations? At the end of the afternoon, we had the sketch of a possible rubric that had been proven to partly work in allowing 10 readers to almost agree on a judgment: that is, we had been (as Broad might have predicted) more successful at generating productive pedagogy-based conversation than choosing a single rubric for assessing the researched essays.

The two of us (Shelley and Sarah) met shortly thereafter to refine the rubric. Although we were concerned about "speaking over" the voices of composition faculty in the name of efficiency, two rubric modifications directly reflected the English 101 faculty voices we had already heard. First came the subdivision of the "Emerging Competence" rating into two categories: "Emerging competence—consistent" and "Emerging competence—inconsistent." While our draft rubric had been designed so that most successful (yet pre-disciplinary) FYC essays would end up with an "Emerging Competence" rating, responses from our rubric workshop participants suggested that instructors needed to separate strong FYC papers from average ones, and might thus use the draft rubric inappropriately if we made no adjustment to the final rubric. We also decided to clarify the rubric's levels in terms of a more recognizable benchmark: students' preparation or lack of preparation for our junior-level required advanced composition course, English 302. The [revised and final rubric](#), though developed specifically by and for English 101 readers, thus connects that course more directly with our WAC program, and, we think, will inform our FYC faculty development in ways that attend to broader disciplinary concerns.

On a Saturday early in spring semester, a group of 10 FYC instructors, 12 including the two of us, met to assess 153 papers for a sample of 10% of the total enrollment of 1553 English 101 students, the lowest recommended percentage, but necessitated by our resource constraints. Only one of the 10 instructors had attended the earlier rubric-generating workshop, so we began both the morning and afternoon sessions by holding "norming" discussions with readers focused on papers written in response to different research assignments that we had—with some difficulty—chosen to exemplify likely features of higher and lower writing competencies. As with the rubric-generating workshop, the discussions about how one feature of an essay might affect its rating compared to another were

encouragingly and sometimes frustratingly rich and wide-ranging. Whereas in most assessment workshops, and particularly in the small department-level WID workshops at Mason, relatively few papers make it to the third-read stage, which occurs when the first two readings don't agree, this assessment ended with 23% of papers needing more than two readings. It is worth noting that our third-read percentage would likely have been even higher had we counted "consistent" and "inconsistent" emergence as separate categories for the purposes of reporting to SCHEV, or had we tried to force very disparate faculty in a de-centered program into the previous three-level rubric. By deciding in advance that collaboration and discussion were at least as valuable to us as a discrete measurement of the achievements of our students and program, as suggested by Elbow and Belanoff, we avoided setting ourselves up for failure. Indeed, participating in the extended process of the creating, splitting, and implementation of the "emerging competence" category—in direct response to local needs and cultures—made us each aware of not just how difficult but also how rich such inclusive, collaborative assessment can be.

### Reflections on the FYC Process

**Sarah:** I felt that the higher level of disagreement among composition faculty versus faculty in the department-based assessment workshops I had helped to facilitate was particularly interesting. The composition instructors seemed to be having trouble separating the creation of this more general rubric from what they themselves were teaching or focusing on in their individual classrooms, whereas discipline-based faculty were able to move more quickly toward establishing generalized expectations regarding how and with which criteria to assess student texts. Seeing first-hand these differences along the continuum of writing at Mason will help me communicate expectations across writing curricula better to all the audiences I address in my liaison role.

**Shelley:** For me, the rubric development and rating processes also illuminated the tension between *inclusive* and *conclusive* assessment methods that I had felt present from the beginning of the process. Initially, as I worked with Ying, Terry, and Sarah, I was pleased to find them receptive to suggestions about where to place English 101 students along the continuum of writing competencies at Mason. As I later watched composition faculty members who were used to grading alone struggle toward agreement in the rubric workshop, I was simultaneously encouraged by the depth of my colleagues' engagement with the process, discouraged by how far apart some of them were from one another and from what the course learning goals and guidelines emphasized, and alarmed at how much time was passing. Each effort to include faculty voices slowed our progress towards concrete results, a challenge that most large programs face. I felt this tension keenly both as an administrator needing to meet deadlines for the assessment results, and as the director of an ostensibly coherent program that was revealed as including an even more diverse set of approaches to and values for teaching student writing than I had anticipated. (Had I been assessing all on my own, or been looking to provide only conclusive results, I might simply have given up at this point.) The conversations with our assessment raters generated the same mix of feelings: the extensive discussions limited the conclusivity of our final results, but brought us into an important and generative discussion, and raised fascinating questions for me about our program, faculty, assignments, and relationship to other writing instruction at Mason, as I discuss below. I remain committed to having our assessment continue as a collaborative, inclusive process. At the same time, having now had the opportunity to collaborate in the design of a mandated assessment program, I can see how setting some clear guidelines and/or limitations can be necessary so that some conclusions can be reached and actions taken to improve the program.

**Ying:** As Shelley and Sarah recounted all the struggles they went through in the implementation stage, I knew they had set another exemplary model for programmatic assessment. In my experience,

faculty's resistance to assessment often comes with the belief that they are already doing assessment in their class, they know what their expectations are for student learning, and they know their students very well. What is often neglected is the consideration of program efficiency and program outcomes—whether the faculty, as a group, are committed to the same goals. I believe the rich conversations at the rubric workshop and the scoring workshop, as Shelley and Sarah described above, are the core of the assessment process. With Sarah taking the coordinating role, the assessment process went smoothly; with Shelley taking the leadership role, "closing the loop" (i.e., using assessment results for curricular improvement) became a seamless part of the teaching and learning process.

## What Each of Us Has Gained from the Planning and Implementation Process

### Ying and OIA

First of all, we have gotten quality data from the assessment process. The results, among other things, will demonstrate how well our students have achieved the learning outcomes for English 101 and how well they may be prepared for writing in their majors. We can evaluate the strengths and weaknesses of our FYC course and identify areas for improvement. The data also allow us to conduct in-depth analyses, such as examining inter-rater reliability and comparing student achievement across different majors and campuses. Our assessment did not stop at the end of data collection and analysis; sharing the results with the faculty and using the results to improve our teaching-with-writing efforts across the curriculum, which Shelley discusses later, is our ultimate outcome. Moreover, the timing of the value-added assessment couldn't be better: Mason's re-affirmation of accreditation by the Southern Association of Colleges and Schools (SACS) is coming up in a year, so the data are being used in that context and also as part of our ongoing academic program review.

### Shelley and the Composition Program

From the perspective of the composition program, our ability to document that nearly a quarter of students' researched-essays were ranked as Not Competent and another third as Inconsistent Emerging Competence—in a program that routinely sends nearly three-fourths of its students forward with A's and B's—provides useful quantitative data. Grade "inflation" is of course a tricky subject to discuss relative to a course taught by new and/or non-tenure-track and/or overworked faculty, and relative to a course that emphasizes writing in multiple stages for multiple genres. Still, the numerical disparity may help provoke useful discussions about program goals and standards.

More immediately beneficial from a faculty development standpoint, however, are the questions raised during the assessment process. While assessing a single research-based essay rather than a portfolio provided a limited snapshot of FYC student writing—and provided challenges in comparing one approach to research-writing to many others—it did focus attention on a genre that has grown substantially in popularity in FYC over the past two decades (see Lunsford and Lunsford). Several faculty readers pointed out that students who did poorly seemed more often to be responding to inappropriate or insufficiently specific assignment prompts: in some cases, assignments designed for maximum student choice and engagement seemed to have been insufficiently scaffolded for less-competent writers. Readers also raised questions about whether a lengthy, wide-ranging, alphabetic "research paper" is in fact appropriate for, or the best use of time in, first-year composition. Students frequently struggled to fill 8-10 pages without resorting to less-credible source use, information dumping, and/or problematic citation strategies. Meanwhile, they demonstrated very tenuous

abilities to synthesize, analyze, and form arguments, core strategies that they will need in other writing assignments in English 302 and in courses across the curriculum.

As a result of composition's participation in this assessment program, we might thus consider developing new assignment structures that meet the English 101 learning goals while also more directly preparing students for other writing tasks they will encounter. Participating in a "value-added" assessment program may not by itself improve teaching and learning in FYC, but framing our rubric in terms of the larger arc of writing instruction and writing learning at Mason has helped us assess English 101 in its dual institutional role: improving students' writing knowledge, skills, and strategies overall, and preparing students for writing tasks elsewhere in the university.

Lastly, the process raised interesting additional questions—as similar experiences have done elsewhere—about the benefits and drawbacks of faculty independence. The most promising of these questions focus less on course content or consistency than on opportunities for faculty growth. Mason faculty who had valued the opportunity to teach and grade independently nonetheless found the normed, anonymous grading protocols of the assessment process intriguing. The experience provoked a steady stream of questions about the roles played in evaluating student writing by such elements as student personality and "improvement," an instructor's idiosyncratic prioritizing of one or another writing strategy (integrating quotations, using topic or thesis sentences, finding credible sources), and the kinds of mid-process interventions (draft-readings, conferences) that instructors created, all of which, of course, influence the grading practices of instructors outside of composition as well. As we move toward another round of collaboration and inclusivity—via resources, workshops, and committee discussions—we hope to extend the benefits of this process well beyond the numbers reported to SCHEV. It becomes clear, overall, that for us "closing the loop" means "opening new discussions" about a range of issues at the core of writing-learning, writing instruction, and the role of FYC in both. Data and participant stories from the assessment process allow us to raise those questions in focused yet inclusive and community-building ways.

### **Sarah as OIA/WAC Liaison**

For me, my role in bringing together the different but related perspectives of faculty and administrators has been invaluable in helping those with a narrower view see the interconnectedness of seemingly discrete processes and requirements. This ability to be multi-voiced, to cross boundaries, and to be a bridge across programs and issues has been rewarding to me personally as well as a strong justification for the creation of a position like mine in the first place.

### **Terry and Writing Across the Curriculum**

The success of the course-embedded FYC assessment reinforces my belief in the WID-assessment model. In addition to the benefits of providing built-in faculty development opportunities, our WID process has also served in many ways as an ideal research site for my scholarship as a WAC professional, which, in turn, informs both my WAC and writing center program development work.<sup>[2]</sup> Personally and professionally, my interactions with faculty across the university have been a source of pleasure, as I share with them the many discoveries they—and I—make about writing in their disciplines and their own individual preferences and predilections. When faculty participants ask me if others have enjoyed talking about student writing and their own teaching practices as much as they have in the workshop or if I have had as much fun leading workshops as I've had leading theirs, I think about what a great job I have. These are the responses that sustain me when I have to persuade resistant faculty and departments that they too must engage in our ongoing assessment of student writing.

## Towards a Conclusion: Our Voices in Unison

While we've each described what we needed and wanted from the assessment process and what each has gained for our programs, in this final section we offer our collective views on how the overall process has fostered the kinds of cross-disciplinary conversations that help us to sustain and enhance our programs. For Shelley, the FYC process and initial results have helped her and the composition faculty to identify conversations they need to have about the role the introductory writing course plays/should play in preparing students for writing in other courses. What are the writing skills and abilities we can expect students to transfer from English 101, to the required advanced WID-focused composition course, and to the writing tasks they encounter across the curriculum? Assessing FYC in the context of discipline-aware advanced composition courses—as well as in relation to Mason's WID major-field courses—reveals some of what makes "transfer" such a complex issue for FYC in particular. What is it that we want to transfer, and to where? Writing "skills," of course, top the list of what faculty and institutions hope will transfer, but it is clear to us that discipline-based analysis strategies, abilities to work in a range of written genres, and a general reflective awareness and resultant flexibility may also be crucial.<sup>[3]</sup>

The "what transfers" question is one that also comes up in the scoring workshops Terry leads, most often in the form of a complaint about why WID faculty still have to teach students to write when they've taken the required composition course(s). For Terry, then, the "emerging competence" final score for the majority of the FYC papers assessed provides additional evidence for the argument that all teachers must take responsibility for helping students develop into fully competent writers in their courses, whether in or outside of the major. This is not a new argument, of course, but one that can now be supported by quantitative data to demonstrate that every course, beyond FYC, must play a role in helping students succeed as writers in college and beyond.

Ying and Sarah have strong evidence that a value-added writing assessment can be carried out successfully based on close collaboration and adequate communication with faculty members. Each has gained a fuller understanding of how to balance the needs and wants of faculty and other program stakeholders with the increasingly frequent assessment mandates they are charged with implementing. Our process has also revealed ways in which assessment might not productively be thought of as "done," the loop "closed," if a collaborative, inclusive mode is to be maintained. As Shelley noted in one of our meetings, she is likely to turn her focus to discussions and changes needed within the composition program once her pre-assessment FYC data is processed—yet it will be productive in the long term for Sarah to be able to poke her head in periodically and draw Shelley back out into the larger culture of writing. And this is what Sarah will continue to do with all of the departments she and Terry have engaged in conversations about their expectations for student writers and writing.

In the end, our faculty may be the ones who have gained the most by participating in these kinds of systematic conversations about student writing with their colleagues, conversations which, in turn, help them to better understand the complex demands placed on student writers not only in their required composition courses but also in their writing-intensive courses across the curriculum. We are confident that the value-added assessment process we have created can and will be sustained. In making sure that all of our voices were heard at the table and in listening to the needs and wants of our own program stakeholders, we've modeled the spirit of negotiation and cooperation that has likewise sustained the culture of writing at Mason.

## References

- Beaufort, Anne. (2007). *College writing and beyond: A New framework for university writing instruction*. Logan, UT: Utah State University Press.
- Broad, Robert. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Logan, UT: Utah State University Press.
- Carter, Michael. (2007). Ways of knowing, doing, and writing in the disciplines. *College Composition & Communication*, 58(3), pp. 385-418.
- Cooper, Charles. (1977). Holistic evaluation of writing. In C. Cooper & L. Odell (Eds), *Evaluating writing: Describing, measuring, judging* (pp. 3-32). Urbana, IL: NCTE.
- Haswell, Richard, & McLeod, Susan. (1997) WAC assessment and internal audiences: A dialogue. In K.Yancey & B. Huot (Eds). *Assessing writing across the curriculum: Diverse approaches and practices* (pp. 217-36). Westport, CT: Ablex.
- Huot, Brian. (2002). *(Re)Articulating writing assessment for teaching and learning*. Logan, UT: Utah State University Press.
- Lunsford, Andrea and Karen Lunsford. (2008) "Mistakes are a fact of life": A national comparative study. *College Composition & Communication*, 59(4), pp. 781-806.
- State Council of Higher Education for Virginia (2007). Guidelines for assessment of student learning. Richmond, VA: SCHEV Task Force on Assessment. (Unpublished report.)
- Sommers, Nancy, & Saltz, Laura. (2004). The Novice as expert: Writing the freshman year. *College Composition & Communication*, 31(1), pp. 124-149.
- Thaiss, Chris, & Zawacki, Terry Myers. (2006). *Engaged writers and dynamic disciplines: Research on the academic writing life*. Portsmouth, NH: Heinemann.
- Thaiss, Chris, & Zawacki Terry Myers . (1997). How portfolios for proficiency help shape a WAC program. In K.Yancey & B. Huot (Eds). *Assessing writing across the curriculum: Diverse approaches and practices* (pp. 79-97). Westport, CT: Ablex.
- Walvoord, Barbara. (2004). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. San Francisco: Jossey-Bass.
- Zawacki, Terry Myers, & Gentemann, Karen. (Forthcoming 2009). Merging a culture of writing with a culture of assessment: Embedded, discipline-based writing assessment. In Paretto, M. & K. Powell (Eds). *Assessment in writing. Assessment in the Disciplines Series* (Vol. 4). Tallahassee: Association of Institutional Research.

## Notes

[1] Just as Richard Haswell and Susan McLeod (1997) suggest about their dialogic approach in "WAC Assessment and Internal Audiences: A Dialogue," we think our "voices at the table" structure models "the ideal spirit" (p. 218) and the collaborative approach by which writing assessment should be carried out.

[2] I've described our course-embedded process and the resulting rubrics, for example, as part of the analysis of the contexts for faculty assignments and expectations that Chris Thaiss and I present in *Engaged Writers and Dynamic Disciplines: Research on the Academic Writing Life*.

[3] See, for example, longitudinal studies that take up questions of transfer, notably and most recently Anne Beaufort's (2007) *College Writing and Beyond* and Nancy Sommers and Laura Saltz' "The Novice as Expert: Writing the Freshman Year."

## Contact Information

Terry Myers Zawacki  
Associate Professor, English

Director, Writing Across the Curriculum Program  
George Mason University  
Telephone: 703-993-1187  
Email: [tzawacki@gmu.edu](mailto:tzawacki@gmu.edu)

E. Shelley Reid  
Associate Professor, English  
Director of Composition  
George Mason University  
Email: [ereid1@gmu.edu](mailto:ereid1@gmu.edu)

Ying Zhou  
Director, Office of Institutional Assessment  
George Mason University  
Email: [yzhou@gmu.edu](mailto:yzhou@gmu.edu)

Sarah E. Baker  
Assistant Director, WAC/Writing Assessment Specialist  
George Mason University  
Email: [sbaker@gmu.edu](mailto:sbaker@gmu.edu)

### Complete APA Citation

Zawacki, Terry Myers, Reid, E. Shelley, Zhou, Ying, & Baker, Sarah E. (2009, December 3). Voices at the table: Balancing the needs and wants of program stakeholders to design a value-added writing assessment plan. [Special issue on Writing Across the Curriculum and Assessment] *Across the Disciplines*, 6. Retrieved from <https://wac.colostate.edu/docs/atd/assessment/zawackietal.pdf>